

Temperature Estimation for a Point of an Infrared Dryer Using Temperature of Neighbouring Points: An Artificial Neural Network Approach

Morteza Mohammadzaheri*, Ali Firoozfar**, Dalileh Mehrabi**, Mohammadreza Emadi** and Abdullah Alqallaf***

*Department of Mechanical and Industrial Engineering, Sultan Qaboos University, Oman

**Islamic Azad University, Iran

***Department of Electrical Engineering, Kuwait University, Kuwait

***Corresponding Author: al.qallaf@ku.edu.kw

ABSTRACT

This paper aims at introduction, design, and validation of a temperature estimation algorithm for an infrared dryer. The proposed algorithm estimates the temperature at one or some points in a thermal system (e.g., an infrared dryer) based on the measured temperature at a number of other points. In this research, the designed algorithm estimates the temperature of a single point; however, the methodology can be evidently extended to multiple points. Inspired by direct and inverse heat transfer models, a mathematical model is presented for this purpose. This model is developed and identified using artificial neural network (ANN) technique and laboratory experimental data. The proposed method exhibits excellent accuracy with no need to thermo-physical properties of the system.

Keywords: Temperature estimation; artificial neural network; infrared dryer; radiation 2

INTRODUCTION

Algorithms to estimate quantities without measurement have been used for several years (Daily et al., 1988, Payton, 1986). At the beginning, they were employed to estimate displacement, velocity, and force in robots and autonomous vehicles (Muir, 1990, Nicholls et al., 1989, Payton, 1986). Then, the application area of these algorithms was extended to estimation of pressure (Corcione et al., 2006), polluting emissions (Liang et al., 2005), wind speed (Kabadayi et al., 2006), and acoustic signals (Halim et al., 2011). Such algorithms have been also presented to estimate temperature (Yu et al., 2014, Mehrabi et al., 2017) mainly using the information of input heat.

In this work, a temperature estimation algorithm is developed for an infrared dryer (an enclosure heated by radiation). This algorithm ideally estimates the temperature distribution on a surface in real-time using the measured temperature of a limited number of points on the surface. In general, if initial conditions, boundary conditions (heat fluxes), geometry, and thermo-physical characteristics of a thermal system are known, and estimation of temperature (distribution) is aimed, the problem is categorised as a direct heat transfer problem. If some temperature information is known and any other factor (e.g., boundary conditions) is unknown, the problem is considered an inverse heat transfer problem (Ozisik, 2000). Inverse heat transfer problems are considered ill-posed and relatively difficult to solve (Tikhonov and Arsenin, 1977). Existing physics-based methods of temperature estimation (based on heat transfer and thermo-dynamics laws) inherit serious difficulties such as temperature-dependency of thermo-physical characteristics (Das, 2015, Baghban et al., 2014) and inevitable large time delay of heat flux estimation (Woodbury and Beck, 2013), which make them impractical for real-time temperature estimation, as further detailed in Results and Discussion section.

In this research, alternatively, a temperature estimation algorithm for irradiative systems has been formulated as a mathematical model in the form of an artificial neural network (Problem Statement section), which is identified using

the experimental data collected in laboratory environment, as detailed in Estimation Algorithm section. Such a model can be supplied together with the final product, i.e., an infrared dryer, to the user; then, the user will have access to the temperature of several points without having sensors on them.

PROBLEM STATEMENT

In general, for any point (i.e., point i) on a surface of an infra-red dryer, e.g., the one depicted in Figure 1, temperature can be estimated in discrete domain as follows (Ghanbari et al., 2010):

$$T_i(k) = F_D \left(T_i(k-1), \dots, T_i(k-r_T), q_{lamp}(k-r_d), q_{lamp}(k-r_d-r_q) \right). \quad (1)$$

where T , q_{lamp} and k are temperature, heat energy emitted by the lamps, and sequence index (a representative of time). r_d , r_T and r_q are delay, temperature, and heat flux orders. F_D is a direct heat transfer model of the dryer (in discrete domain) and includes its thermo-physical and geometrical characteristics. A detailed thermal analysis and modelling of the dryer are presented in Appendix A. Physics-based direct heat transfer models may be used to produce data as alternatives to experimental setups of thermal systems to avoid experimentation difficulties (Kowsary et al., 2006). In addition to physics-based models, data-driven artificial neural network (ANN) models have been also presented as direct heat transfer models of infrared dryers (Ghanbari et al., 2010).

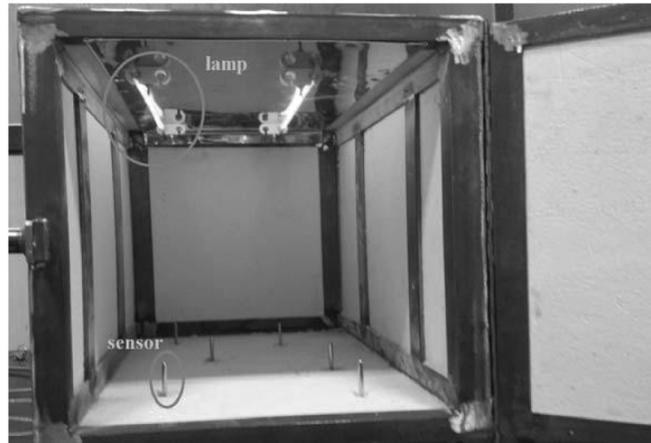


Fig. 1. The investigated dryer.

Besides the problem of direct heat transfer modelling (or in short, direct modelling), an inverse problem for heat flux estimation has been also formulated and solved in discrete domain (Mirsephai et al., 2012):

$$q_{lamp}(k) = F_I \left(T_j(k+r_d), \dots, T_j(k+r_T+r_d) \right). \quad (2)$$

where F_I is an inverse heat transfer model, which estimates the emitted energy of the heat source using the temperature of any point (with a temperature sensor, i.e., point j) on a surface of the infra-red dryer. For the purpose of real-time temperature estimation without use of emitted heat information, a combination of Equations (1) and (2) can be roughly presented in Equation (3). This equation removes the role of emitted heat and relates the temperature of different points (j and i). However, Equation (3) is approximate as the influence of future values of temperature on this model has been disregarded due to their unavailability in real-time.

$$T_i = F \left(T_j, \text{history of } T_j, \text{history of } T_i \right). \quad (3)$$

History of temperatures represent the effect of emitted heat on the temperature at sequential instants. With removing the emitted heat, modelling of this cause-effect relationship is not a major concern. Therefore, histories are proposed to be removed from Equation (3). Instead, the measured temperature at a number of points (with sensors) is added to the formula or

$$T_i = F_S(T_j, j \in [1, n], j \neq i). \quad (4)$$

where F_S is an estimation formula, which receives the information from a limited number ($n-1$) of temperature sensors on a surface of an enclosure heated mainly by radiation, then accurately estimates the temperature of another point on the surface. In the next session, a method for developing F_S is presented and its applicability is experimentally shown in this research for a sample dryer depicted in Figure 1. There exist six thermocouples in the dryer ($n=6$) and the temperature of the last one is chosen to be estimated using the measured temperatures of others, or $i=6$ in Equation (4). Validation of the estimation formula for a single point shows the feasibility of the idea and demonstrates the possibility of designing more comprehensive multi-output formulae/algorithms to estimate temperature at multiple points. Section 3 reports the development and identification of F_S (the estimation formula) for the proposed case study.

DEVELOPMENT OF ESTIMATION ALGORITHM

Experimental Setup, Data Collection and Preparation

Figure 2 shows the experimental setup. The dryer has the dimensions of 60cm×40cm×40cm. The heat sources are two 39N3 Toshiba 1000W infrared halogen lamps. Asymmetrically arranged K-type thermocouples, equipped by AD595 Analog Device ICs as signal amplifiers, measure the temperature. The walls of the dryer are made of insulation board with the thickness of 25mm and mounted on a steel skeleton built with 2mm thick steel plates.

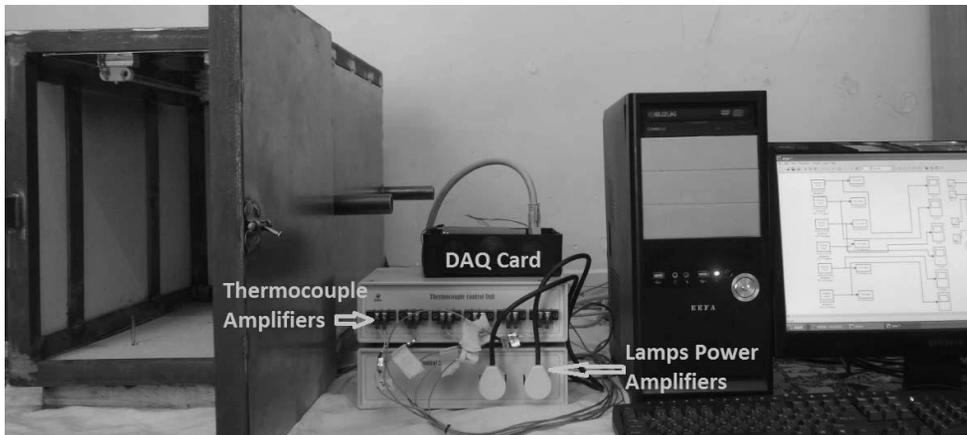


Fig. 2. The main components of the experimental setup.

For data collection, MathWorks Simulink software and its Real-Time Workshop and Real-Time Windows Target (RTWT) toolboxes were employed. This software platform produces executive codes to connect hardware (the thermocouples and the lamps) to the computer through an Advantech PCI-1710U data acquisition (DAQ) card. A power amplifier was also designed, built, and used to magnify the signals coming from the computer to the lamp with a circuit depicted in Appendix B.

18 different sets of experiments were carried out. Different triangular functions of power with peaks of 400, 600, and 800 W and the time periods of 100s, 120s, 140s, 160s, 180s, and 200s were applied on the lamps. Each experiment took five minutes, but the data collected at the first two seconds were disregarded. The data measured by

all thermocouples were collected at the sampling frequency of 10 Hz (3000 samples at each experiment). The data of 17 experiments were used for modelling and validation, and the data of one set with a peak power of 600W per lamp and time period of 160s were solely used to test the model as detailed in Result and Discussion section.

The whole collected data pass through a low pass filter with a cut-off frequency of 2 Hz to reduce the noise. The filtered data then resampled up to the frequency of 1Hz to match fairly slow dynamics of the dryer (Mohammadzaheri et al., 2015) and mapped into the range of [-1 1] before training (as a type of normalisation) (Demuth et al., 2008). The effect of normalisation on the model is detailed in Appendix C.

ANN Design and Training

An ANN is proposed to play the role of F_s in Equation (4), i.e., the estimation formula. The proposed ANN model is a fully connected multi-layer-perceptron (Mohammadzaheri et al., 2012). Considering Equation (4) and $n=i=6$, the ANN has five inputs and one output. One hidden layer with sigmoid activation functions of f , presented in Equation (5), has been proved to be sufficient to make this ANN a universal approximator (Chen et al., 1995):

$$f(x) = \frac{2}{1 + \exp(-2x)} - 1. \tag{5}$$

The recommended number of neurons in such a single hidden layer is two times the number of inputs plus one (Haykin, 1999), i.e., 11. Figure 3 and Equation (6) illustrate the proposed ANN:

$$T_6 = \sum_{i=1}^{11} P_i f \left(\sum_{j=1}^5 W_{ij} T_j + B_i \right) + b. \tag{6}$$

where W_{ij} and P_j are the first and the second layer weights and B_j and b are biases.

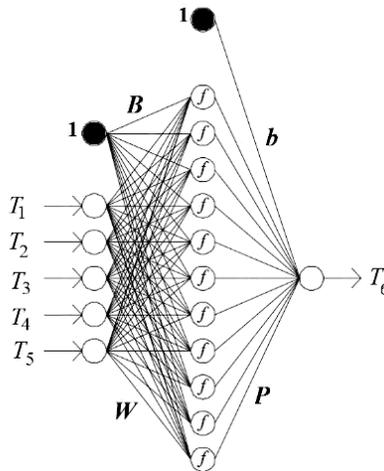


Fig. 3. A schematic of the employed artificial neural network.

Discrepancy of the output of a model (e.g., ANN) and the measured output (i.e., T_6) with same inputs is represented by the ‘model error’. Tuning the parameters of an ANN to minimise the model error for so called training data is named ‘training’. Back-propagation batch training with Levenberg-Marquardt method, a recursive algorithm detailed in Appendix C, was employed for training in this research in conjunction with Nguyen-Widrow weights initialisation algorithm (Nguyen and Widrow, 1990a). 4560 randomly chosen samples of data amongst 17 sets of experimental data mentioned in subsection 3-1 were employed for training and the rest for validation. That is, at each iteration of ANN

training, the model error of the validation data was observed as well as the training error; a discrepancy in trend of these dual errors is considered as a sign of over-fitting (over-training). Over-fitting decreases the inclusiveness of the ANN (Mohammadzaheri et al., 2007). In the case of over-fitting, training should be stopped. The trained and validated ANN was then tested by a single series of the data, which were used for neither training nor validation; therefore, the requirements of cross-validation were fulfilled.

RESULTS AND DISCUSSIONS

Figure 4 shows the test result. The mean of absolute error is $0.425\text{ }^{\circ}\text{C}$, which is excellent. It should be noted that the temperature data have been filtered before ANN training; as a result, the ANN simulates filter dynamics as well. The employed type of thermocouples has noise in the range of $[-2\ 2]\text{ }^{\circ}\text{C}$ or even higher (Reotemp-Instruments, 2015).

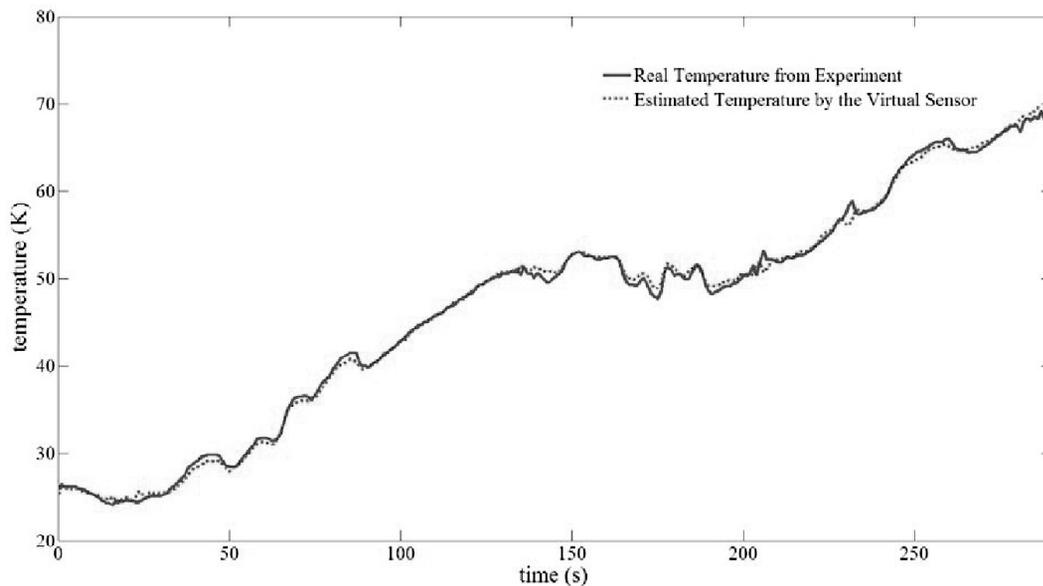


Fig. 4. Test result for estimation.

A physics-based heat transfer algorithm might be considered for temperature estimation either. However, this approach would need geometrical and thermo-physical characteristics of the dryer, which could be very difficult to estimate, particularly with considering the fact that thermo-physical characteristics are temperature-dependent (Baghban et al., 2014). Even with exact knowledge of characteristics (e.g., in simulation), the accuracy and the time delay of estimation based on existing physics-based methods would be a concern. For instance, inaccuracy of heat flux estimation, which is essential for such estimation methods, is an issue. Many of heat flux estimation problems are solved for the entire operation time to increase the estimation accuracy (Kameli and Kowsary, 2014); these the so called whole-domain solutions are not applicable for real-time temperature estimation. Even for less accurate sequential (non-whole-domain) heat flux estimation, a significant time delay of estimation is expected (Kowsary et al., 2006). In fact, with currently available methods of physics-based inverse heat transfer, developing accurate real-time temperature estimation algorithms (based on temperature of other points) seems to be impossible, particularly for complex irradiative thermal systems. On the other hand, as a prominent advantage, the proposed method presents outstanding accuracy with no need to thermo-physical information of the dryer.

As potential disadvantages for use of artificial intelligence, (i) the proposed method needs experimental data even at target points for estimation, and (ii) in the case of any significant change in the operating condition (e.g., having different materials within the dryer), new sets of data need to be collected and used to develop a new ANN

model. However, (i) data collection process needs to be performed only once on a single dryer in the laboratory of the manufacturer, and the final user can utilise the developed algorithm for temperature estimation with no need to data collection. (ii) Also, a number of models may be developed for a variety of prevalent operating conditions to be employed as appropriate by the user.

CONCLUSION

A real-time temperature estimation algorithm was proposed, designed, and validated for a sample infrared dryer. This algorithm was aimed to estimate the temperature at a point on a surface of the dryer using the measured temperature of a number of other points. First of all, considering direct and inverse heat transfer models of the system, the problem was formulated as a mathematical modelling task. Then a universal-approximator ANN was opted as the model.

Afterwards, experimental data were collected, filtered, re-sampled, arranged, and used to train and cross-validate the model. The developed algorithm exhibits excellent test accuracy with no need to any information about the geometry and more importantly thermo-physical properties of the dryer, which may be difficult to estimate. The algorithm can be straightforwardly extended to estimate the temperature of other points through further measurements in laboratory environment and developing multi-output ANNs.

REFERENCES

- Baghban, M., Mansouri, S.H. & Shams, Z. 2014.** Inverse radiation-conduction estimation of temperature-dependent emissivity using a combined method of genetic algorithm and conjugate gradient. *Journal of Mechanical Science and Technology*, **28**: 739-745.
- Chen, T.P., Chen, H. & Liu, R.W. 1995.** Approximation capability in C(R,N) by multilayer feedforward networks and related problems. *IEEE Transactions on Neural Networks*, **6**: 25-30.
- Corcione, G.E., Lavorgna, M., Palma, G. & Scognamiglio, O. 2006.** Virtual pressure sensor for a common rail injection system. Google Patents.
- Daily, M.J., Harris, J.G. & Reiser, K. 1988.** An operational perception system for cross-country navigation. in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Das, R. 2015.** Estimation of parameters in a fin with temperature-dependent thermal conductivity and radiation. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*: 0954408915575386.
- Demuth, H., Beale, M. & Hagan, M. 2008.** *Neural network toolbox™ 6. User's guide*.
- Erturk, H., Ezekoye, O.A. & Howell, J.R. 2002.** The application of an inverse formulation in the design of boundary conditions for transient radiating enclosures. *Journal of Heat Transfer-Transactions of the Asme*, **124**: 1095-1102.
- Ghanbari, M., Mirsepahi, A., Mohammadzaheri, M., Abhary, K. & Chen, L. 2010.** Neural Network Based Solution for Modelling of an Infrared Furnace. in: *Chemeca, Engineering at the Edge, Adelaide, South Australia*.
- Halim, D., Cheng, L. & Su, Z. 2011.** Virtual sensors for active noise control in acoustic- structural coupled enclosures using structural sensing: robust virtual sensor design. *The Journal of the Acoustical Society of America*, **129**: 1390-1399.
- Haykin, S. 1999.** *Neural Networks A Comprehensive Introduction*, Prentice Hall, New Jersey.
- Jang, J.R., Sun, C. & Mizutani, E. 2006.** *Neuro-Fuzzy and Soft Computing*, New Delhi, Prentice-Hall of India.
- Kabadayi, S., Pridgen, A. & Julien, C. 2006.** Virtual sensors: Abstracting data from physical sensors. in: *International Symposium on World of Wireless, Mobile and Multimedia Networks*.
- Kameli, H. & Kowsary, F. 2014.** A new inverse method based on Lattice Boltzmann method for 1D heat flux estimation. *International Communications in Heat and Mass Transfer*, **50**: 1-7.
- Kowsary, F., Mohammadzaheri, M. & Irano, S. 2006.** Training based, moving digital filter method for real time heat flux function estimation. *International communications in heat and mass transfer*, **33**: 1291-1298.
- Liang, C.Y., Srinivasan, S. & Jacobson, E.E. 2005.** NOx emission-control system using a virtual sensor. Google Patents.

- Mehrabi, D., Mohammadzaheri, M., Firoozfar, A. & Emadi, M. 2017.** A fuzzy virtual temperature sensor for an irradiative enclosure. *Journal of Mechanical Science and Technology*, **31**: 4989-4994.
- Mirsepahi, A., Mohammadzaheri, M., Chen, L. & O'neill, B. 2012.** An artificial intelligence approach to inverse heat transfer modeling of an irradiative dryer. *International Communications in Heat and Mass Transfer*, **39**: 40-45.
- Mohammadzaheri, M. & Chen, L. 2008.** Intelligent Modelling of MIMO Nonlinear Dynamic Process Plants for Predictive Control Purposes. in: *The 17th World Congress of The International Federation of Automatic Control*, Seoul, Korea.
- Mohammadzaheri, M. & Chen, L. 2010.** Intelligent Predictive Control of Model Helicopters' Yaw Angle. *Asian Journal of Control*, **12**: 1-13.
- Mohammadzaheri, M., Chen, L. & Grainger, S. 2012.** A critical review of the most popular types of neuro control. *Asian Journal of Control*, **16**: 1-11.
- Mohammadzaheri, M., Chen, L., Mirsepahi, A., Ghanbari, M. & Tafreshi, R. 2015.** Neuro Predictive Control of an Infrared Dryer with a Feedforward Feedback Approach. *Asian Journal of Control*, **17**: 1972-1977.
- Mohammadzaheri, M., Mirsepahi, A., Asef-Afshar, O. & Koohi, H. 2007.** Neuro-fuzzy modeling of superheating system of a steam power plant. *Applied Math. Sci*, **1**: 2091-2099.
- Mohammadzaheri, M., Tafreshi, R., Khan, Z., Franchek, M. & Grigoriadis, K. 2016.** An intelligent approach to optimize multiphase subsea oil fields lifted by electrical submersible pumps. *Journal of Computational Science*, **15**: 50-59.
- Muir, P. F. 1990.** A virtual sensor approach to robot kinematic identification: theory and experimental implementation. in: *IEEE International Conference on Systems Engineering*.
- Nguyen, D. & Widrow, B. 1990a.** Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. in: *International Joint Conference on Neural Networks*.
- Nguyen, D. & Widrow, B. 1990b.** Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. in: *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*.
- Nicholls, H., Rowland, J. & Sharp, K. 1989.** Virtual devices and intelligent gripper control in robotics. *Robotica*, **7**: 199-204.
- Ozisik, M.N. 2000.** *Inverse heat transfer: fundamentals and applications*, CRC Press.
- Payton, D. W. 1986.** An architecture for reflexive autonomous vehicle control. in: *IEEE International Conference on Robotics and Automation*.
- Reotemp-Instruments. 2015.** *Thermocouple Info* [Online]. Available: <http://www.thermocoupleinfo.com/> [2015].
- Tikhonov, A.N. & Arsenin, V.I.a.K. 1977.** *Solutions of ill-posed problems*, Vh Winston.
- Woodbury, K.A. & Beck, J.V. 2013.** Estimation metrics and optimal regularization in a Tikhonov digital filter for the inverse heat conduction problem. *International Journal of Heat and Mass Transfer*, **62**: 31-39.
- Yu, Y., Woradechjumroen, D. & Yu, D. 2014.** Virtual surface temperature sensor for multi-zone commercial buildings. *Energy Procedia*, **61**: 21-24.

Submitted: 22/11/2016

Revised: 08/02/2018

Accepted: 11/02/2018

APPENDIX A -DRYER DYNAMICS

A mathematical model of the dryer, depicted in Figure.1, can be developed based on the first law of thermodynamics (Mirsephai et al., 2012). The dryer (with closed doors) can be assumed as an enclosure with Diffuse-Grey Surfaces. The following assumptions are considered for thermo-dynamics-based modelling (Erturk et al., 2002, Mirsephai et al., 2012):

1. The surface properties are non-uniform.
2. ϵ , the emissivity factor of the surfaces, is independent of the wavelength and the direction of radiation.
3. All the energy is emitted and reflected diffusely.

Incident and reflected energy flux is non-uniform; as a result, in order to find the mathematical model of the dryer, its surfaces to be divided into infinitesimal areas. A tiny element on a surface (Figure (A.1)) is investigated as following:

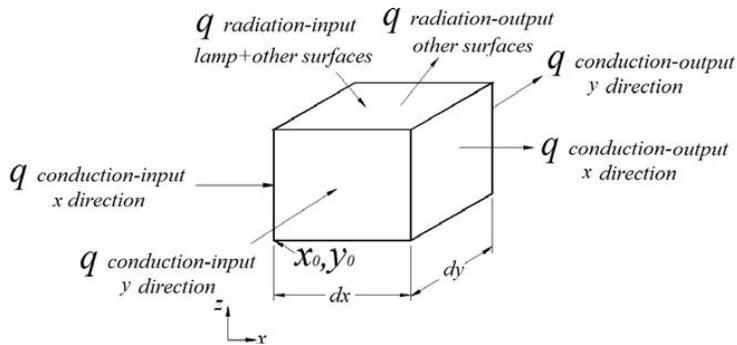


Fig. A.1. An element on bottom internal surface of the furnace.

In general, in terms of energy:

$$E_{Input} - E_{Output} + E_{Generation} - E_{Consumption} = E_{Accumulation} \tag{A.1}$$

The input energy to this element comes from the conduction in x and y directions, the radiation from the lamps, and the radiation from other surfaces:

$$E_{Input} = -k_{cond} \frac{\partial T(x, y, t)}{\partial x} \Big|_{x=x_0} l_z dy - k_{cond} \frac{\partial T(x, y, t)}{\partial y} \Big|_{y=y_0} l_z dx + q_{lamps}(x, y, t) + \sum_{j=1}^N \epsilon_j \sigma T_j^4 F_j(x, y) \Big|_{(x,y)=(x_0,y_0)} dx dy \tag{A.2}$$

where k_{cond} is the heat conduction coefficient (W/mK), t is time (s), l_z is the thickness of the dryer body (m), T and T_j are exact and average temperature (K), q is heat flux (W), j is the index of elements on the surfaces rather than the modelled one, and N is the number of these elements. $F_j(x, y)$ is the exchange factor of the inner surface of j^{th} element and the element located at (x, y) .

The output energy can be divided in two categories: output energy through conduction, radiation due to high temperature of the element and the reflecting radiation:

$$E_{Output} = -k_{cond} \left. \frac{\partial T(x, y, t)}{\partial x} \right|_{x=x_0+dx} l_z dy - k_{cond} \left. \frac{\partial T(x, y, t)}{\partial y} \right|_{y=y_0+dy} l_z dx + \epsilon \sigma T^{-4} dx dy, \quad (A.3)$$

where σ is the Stefan-Boltzmann Constant = $5.6703 \times 10^{-8} \text{ W/m}^2 \text{ K}^4$, ϵ is the emissivity factor of the element and \bar{T} is the average temperature of the inner side of the element.

According to Taylor's series:

$$\left. \frac{\partial T(x, y, t)}{\partial x} \right|_{x=x_0+dx} \cong \left. \frac{\partial T(x, y, t)}{\partial x} \right|_{x=x_0} + \left. \frac{\partial^2 T(x, y, t)}{\partial x^2} \right|_{x=x_0} dx, \quad (A.4)$$

and

$$\left. \frac{\partial T(x, y, t)}{\partial y} \right|_{y=y_0+dy} \cong \left. \frac{\partial T(x, y, t)}{\partial y} \right|_{y=y_0} + \left. \frac{\partial^2 T(x, y, t)}{\partial y^2} \right|_{y=y_0} dy. \quad (A.5)$$

also

$$E_{Generation} = E_{Consumption} = 0. \quad (A.6)$$

$$E_{Accumulation} = \rho V C_p \frac{\partial T(x, y, t)}{\partial t} = \rho (l_z dx dy) C_p \frac{\partial T(x, y, t)}{\partial t}, \quad (A.7)$$

where V is the volume of the element, C_p is the specific heat capacity at constant pressure and ρ is the density of element. After considering (A.2-7) in (A.1) and dividing all terms by V ,

$$\left. k_{cond} \frac{\partial^2 T(x, y, t)}{\partial x^2} \right|_{x=x_0} + \left. k_{cond} \frac{\partial^2 T(x, y, t)}{\partial y^2} \right|_{y=y_0} - \frac{\epsilon}{l_z} \sigma T^{-4} + \frac{1}{l_z} \sum_{j=1}^N \sigma T_j^{-4} F_j(x, y) \Big|_{(x,y)=(x_0, y_0)} + \frac{q_{lamp}(x, y, t)}{V \partial t} = \rho C_p \partial T(x, y, t). \quad (A.8)$$

Due to insulation, the following boundary conditions can be considered for the dryer (see Figure A.2):

$$T(x, y, 0) = T_0, \quad (A.9)$$

$$\frac{\partial T(l_x, y, t)}{\partial x} = 0, \quad (A.10)$$

$$\frac{\partial T(x, l_y, t)}{\partial y} = 0. \quad (A.11)$$

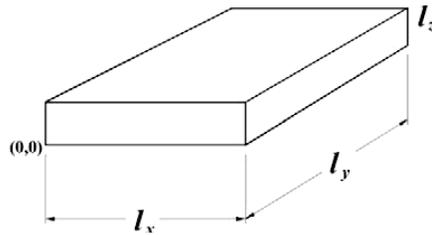


Fig. A.2. A wall of the dryer.

Simultaneous solution of Equation (A.8) with boundary conditions of Equations (A.9-11) for all elements on the surfaces of the dryer returns temperature distribution on those surfaces.

APPENDIX B- LAMPS POWER AMPLIFIER CIRCUIT

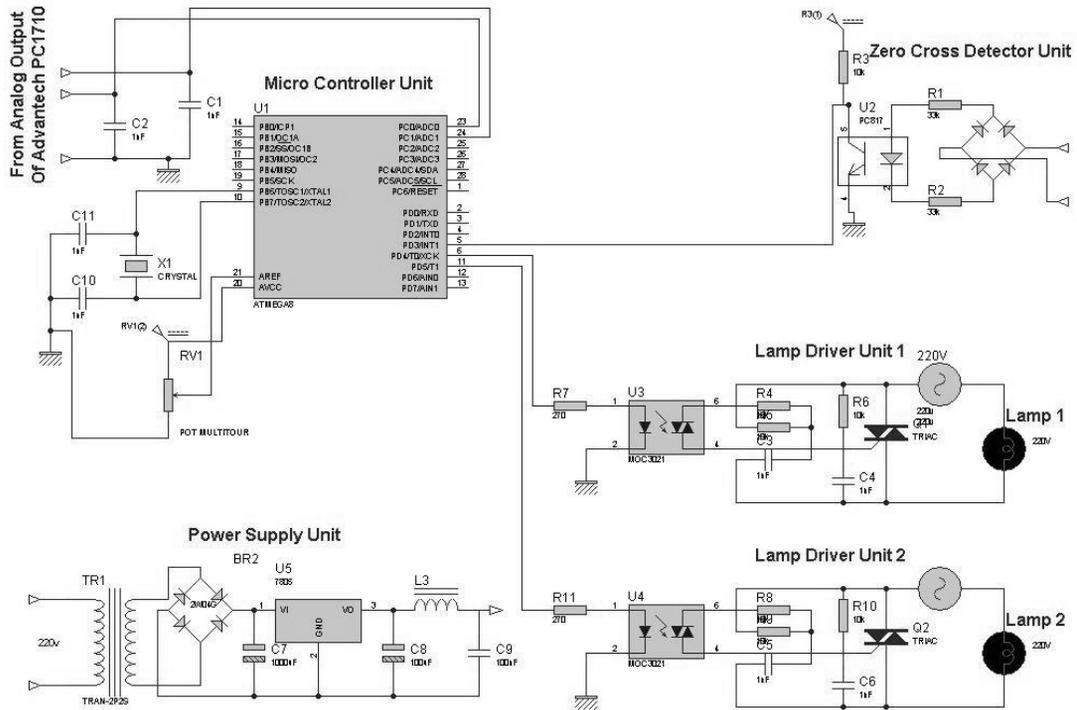


Fig. B.1. Lamps Power Amplifier.

APPENDIX C- NORMALISATION AND ANN TRAINING

Before training ANNs, the input and output data columns are often normalised to remove the effect of magnitude difference between variables (Mohammadzaheri and Chen, 2008). A common way of normalisation, employed in this research, is to map the data from their real range into the range of [-1 1]. The input(s) to an ANN, trained by the normalised data, should be normalised and the output(s) should be de-normalised (de-mapped) into the real range. Normalisation and de-normalisation stages may be embedded into the final presentation of the ANN to remove the need for normalisation/de-normalisation of input(s)/output(s).

After normalisation, the first stage to train an ANN is to define the model error, which represents the discrepancy between the output of the ANN and the real system. A popular definition of the model error, and the one used in this research, is mean of squared errors, which was presented as Equation (C.1) or Equation (C.2):

$$E = \left(y_{\text{ANN}} - y_{\text{system}} \right)^2, \tag{C.1}$$

$$E = \sum_{i=1}^n \left(y_{\text{ANN}} - y_{\text{system}} \right)^2, \tag{C.2}$$

where y represents the output(s) of the system or the ANN, and n_d is the number of training data sets. Equations (C.1) and (C.2) are used when a single set of input-output data or the whole data is used for error calculation, respectively, i.e., single pattern or batch training.

y_{ANN} and consequently the error function are influenced by the weights and the biases shown in Equation (6), i.e., 78 parameters in the proposed ANN. Thus, E is presented as $E(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of all the parameters of the ANN. Training is the process of tuning $\boldsymbol{\theta}$ components so as to minimise (optimise) $E(\boldsymbol{\theta})$. An approach to tackle this optimisation problem is to approximate the model error using the Taylor series up to the second order derivatives:

$$E(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \cong E(\boldsymbol{\theta}) + \frac{\partial E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (\Delta\boldsymbol{\theta}) + \frac{1}{2} \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} (\Delta\boldsymbol{\theta})^2. \tag{C.3}$$

The goal is to find $\boldsymbol{\theta}$ so that the error converges towards its minimum value. Since $\boldsymbol{\theta}$ is a vector rather than a scalar, the derivatives are in the following form:

$$\frac{\partial E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{g} = \left[\frac{\partial E(\boldsymbol{\theta})}{\partial \theta_1}, \dots, \frac{\partial E(\boldsymbol{\theta})}{\partial \theta_n} \right]^T$$

$$\frac{\partial^2 E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} = \mathbf{H} = \begin{bmatrix} \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1^2} & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} & \dots & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_{n_p}} \\ \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_2^2} & \dots & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_{n_p}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_{n_p-1}} & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_{n_p-1}} & \dots & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_{n_p-1}^2} \\ \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_{n_p}} & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_{n_p}} & \dots & \frac{\partial^2 E(\boldsymbol{\theta})}{\partial \theta_{n_p}^2} \end{bmatrix},$$

where n_p is the number of parameters. A solution to this optimisation problem is presented in Equation (C.4), namely, Newton direction (Jang et al., 2006):

$$\Delta\boldsymbol{\theta} = -\mathbf{H}^{-1} \mathbf{g}. \tag{C.4}$$

However, Equation (C.4) is applicable only if H is invertible. Levenberg and Marquardt (Jang et al., 2006) presented an alternative to improve and generalize Equation (C.4):

$$\Delta\boldsymbol{\theta} = -\eta (\mathbf{H} + \lambda \mathbf{I})^{-1} \mathbf{g}. \tag{C.5}$$

where \mathbf{I} is the unit matrix with the size of n_p , and λ is the smallest number that can make the matrix within the parenthesis invertible; η is calculated through linear search. Algorithms to find η and λ have been detailed in Mohammadzaheri and Chen (2010) and Jang et al. (2006). As a prerequisite to calculate \mathbf{H} and \mathbf{g} , E and its derivatives are analytically presented as functions of $\boldsymbol{\theta}$, namely, error back-propagation.

However, as a drawback of all derivative based optimisation algorithms including Levenberg-Marquardt, the algorithm may be trapped in a local minimum. That is, training results in ANN parameters, which do not lead to the minimum model error. If training is restarted from the same initial values of parameters, the algorithm moves towards the same trap again. As a result, an initialisation algorithm is essential to assign appropriate initial values to the parameters of the ANN at the beginning or after each unsuccessful training. Such an algorithm should have two features: (i) The assigned initial values are not very far from the best values of parameters to reduce both computational effort and the chance of being trapped in local minima; this demands the use of training data. (ii) Randomness should be included to make sure initial conditions leading to local minima are not repeated. Nguyen and Widrow presented such an algorithm in 1990, which is still widely used (Nguyen and Widrow, 1990a).

APPENDIX D- AN INTRODUCTION TO NGUYEN-WIDROW ALGORITHM

The ANNs employed in this research are universal approximators and have sigmoid activation functions in their hidden layers. As shown in Fig.D.1, a sigmoid function is nearly linear for an interval.

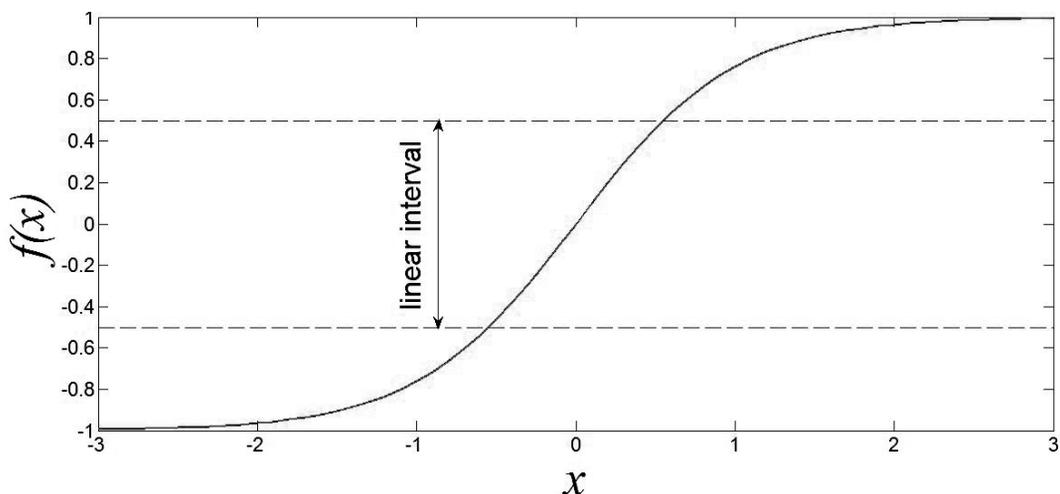


Fig. D.1. The sigmoid function presented by (6) and its linear interval (Mohammadzaheri et al., 2016).

In practice, if an input to a sigmoid activation function is outside the range associated with the linear interval, it will have a slight effect on the output; thus, its effect on the error and training process will be insignificant. Such a situation slows training down. For given input values, ANN weights determine if the output of a sigmoid function is located within linear interval. Nguyen-Widrow algorithm suggests initial values for the weights, considering the range of inputs, so that the output of the activation function lies within the linear interval of the sigmoid function. Such an deliberate initialisation accelerates the training process. Random functions are also used in this algorithms to avoid repeated initial weights (Nguyen and Widrow, 1990b).

تقدير درجة الحرارة لنقطة في مجفف الأشعة تحت الحمراء باستخدام درجة الحرارة من النقاط المجاورة، نهج الشبكة العصبية الاصطناعية

*مرتضى محمد ظاهري، **علي فيروزفر، **دليله مهرايبي، *محمد رضا عمادي و***عبد الله القلاف
*قسم الهندسة الميكانيكية والصناعية، جامعة السلطان قابوس، عمان
**جامعة آزاد الإسلامية، إيران
***قسم الهندسة الكهربائية، جامعة الكويت، الكويت

الخلاصة

يهدف هذا البحث إلى تعريف وتصميم والتحقق من خوارزمية تقدير درجة الحرارة لمجفف الأشعة تحت الحمراء. تقوم الخوارزمية المقترحة بتقدير درجة الحرارة عند نقطة واحدة أو بعض النقاط في نظام حراري (مثل المجفف بالأشعة تحت الحمراء) استناداً إلى درجة الحرارة المقاسة في عدد من النقاط الأخرى. في هذا البحث، تقوم الخوارزمية بتقدير درجة حرارة نقطة واحدة؛ ومع ذلك، يمكن توسيع المنهجية المقترحة بوضوح إلى نقاط متعددة. تم تقديم نموذج رياضي لهذا الغرض مستوحى من نماذج نقل الحرارة المباشرة والعكسية. تم تطوير هذا النموذج وتحديد استخدامه تقنية الشبكة العصبية الاصطناعية (ANN) والبيانات التجريبية المختبرية. تُظهر الطريقة المقترحة دقة ممتازة دون الحاجة إلى الخصائص الفيزيائية الحرارية للنظام.